Is automatic or manual transmission better for mileage?

Motor Trend, a US car magazine, is interested in exploring the relationship between the type of transmission (automatic or manual) and fuel consumption/mileage using data they have collected on different US car models in 1973-74. Our exploratory data analysis reveals that, without accounting for other characteristics of automobile design and performance, cars with automatic transmission have significantly worst mileage (-7.24 mpg, 95% confidence interval: -11.28 -3.21 mpg) than cars with manual transmission.

However, our analysis shows that this effect is mainly due to the fact that automatic cars in the seventies were equipped with bigger engines (and thus were also heavier) than manual cars. When accounting for this, the difference in mileage between automatic and manual transmission is in fact statistically insignificant.

This is likely to reflect the technological limitations of automatic transmission design at that time, since these old automatic cars had fewer forward gears and smaller rear axle ratios than manual cars (or even modern automatic cars) and thus required more torque to achieve similar performance/acceleration.

Exploratory data analysis and regression modeling

```
data(mtcars)
head(mtcars, 3)
```

```
##                mpg cyl disp  hp drat    wt  qsec vs am gear carb
## Mazda RX4     21.0   6  160 110 3.90 2.620 16.46  0  1    4    4
## Mazda RX4 Wag 21.0   6  160 110 3.90 2.875 17.02  0  1    4    4
## Datsun 710    22.8   4  108  93 3.85 2.320 18.61  1  1    4    1
```

The dataset (mtcars) includes fuel consumption in miles per gallon (mpg) and 10 characteristics of automobile design and performance for 32 different US car models in 1973-74. These characteristics are: number of cylinders (cyl), displacement (disp, in cubic inches), horsepower (hp), rear axle ratio (drat), weight (wt, in thousands of pounds), acceleration (qsec, 1/4 mile time in seconds), configuration of cylinders (vs, 0 = v-shaped, 1 = straight/in-line), type of transmission (am, 0 = automatic, 1 = manual), number of forward gears (gear), number of carburetors (carb).

As already mentioned in the executive summary above, cars with automatic transmission have significantly worst mileage than cars with manual transmission (see sample estimates and 95% confidence interval below).

```
t.test(mpg ~ am, data = mtcars)
```

```
##
##  Welch Two Sample t-test
##
## data:  mpg by am
## t = -3.7671, df = 18.332, p-value = 0.001374
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -11.280194  -3.209684
## sample estimates:
##   mean in group auto mean in group manual
##             17.14737             24.39231
```

```r
fit = lm(mpg ~ am, data = mtcars)
summary(fit)
```

```
##
## Call:
## lm(formula = mpg ~ am, data = mtcars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -9.3923 -3.0923 -0.2974  3.2439  9.5077
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   17.147      1.125  15.247 1.13e-15 ***
## ammanual       7.245      1.764   4.106 0.000285 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.902 on 30 degrees of freedom
## Multiple R-squared:  0.3598, Adjusted R-squared:  0.3385
## F-statistic: 16.86 on 1 and 30 DF,  p-value: 0.000285
```

However, when accounting for the fact that automatic cars in the seventies required bigger engines to achieve similar performance/acceleration (see Appendix), the difference in mileage between automatic and manual transmission is no longer statistically significant.

```r
fit2 = lm(mpg ~ am + disp, data = mtcars)
summary(fit2)
```

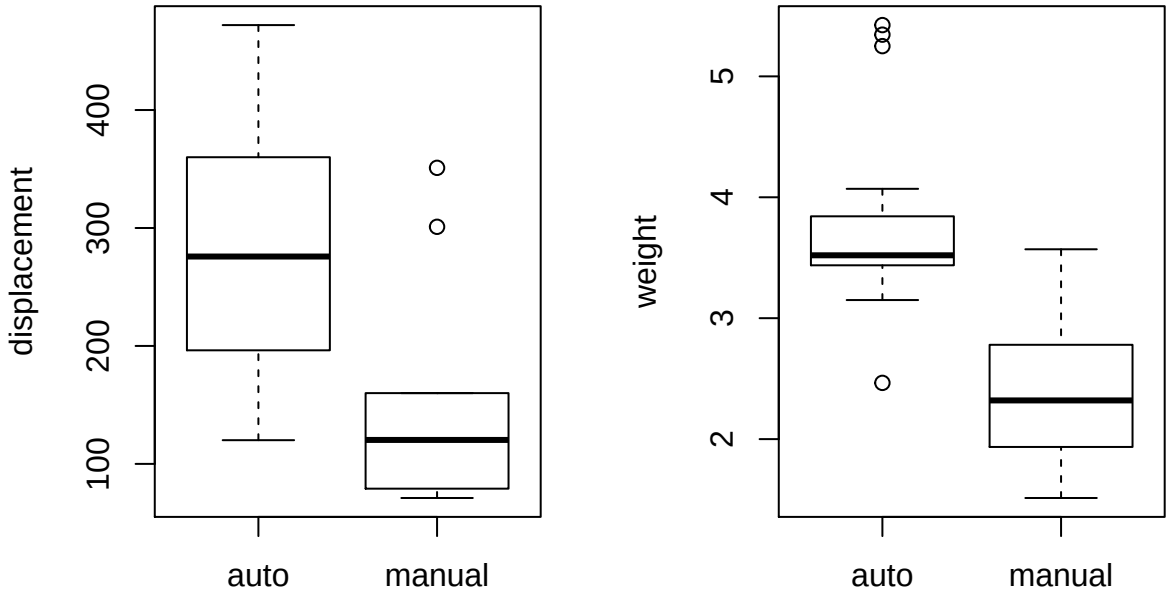
```
##
## Call:
## lm(formula = mpg ~ am + disp, data = mtcars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.6382 -2.4751 -0.5631  2.2333  6.8386
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 27.848081   1.834071  15.184 2.45e-15 ***
## ammanual     1.833458   1.436100   1.277    0.212
## disp        -0.036851   0.005782  -6.373 5.75e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.218 on 29 degrees of freedom
## Multiple R-squared:  0.7333, Adjusted R-squared:  0.7149
## F-statistic: 39.87 on 2 and 29 DF,  p-value: 4.749e-09
```

Appendix

As already mentioned in the executive summary, automatic cars in the seventies were equipped with bigger engines (and thus were also heavier) than manual cars.

```r
par(mfrow=c(1,2))

boxplot(disp ~ am, data = mtcars, ylab="displacement")
boxplot(wt ~ am, data = mtcars, ylab="weight")
```




```r
t.test(disp ~ am, data = mtcars)
```

```
##
##  Welch Two Sample t-test
##
## data:  disp by am
## t = 4.1977, df = 29.258, p-value = 0.00023
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   75.32779 218.36857
## sample estimates:
##   mean in group auto mean in group manual
##             290.3789             143.5308
```

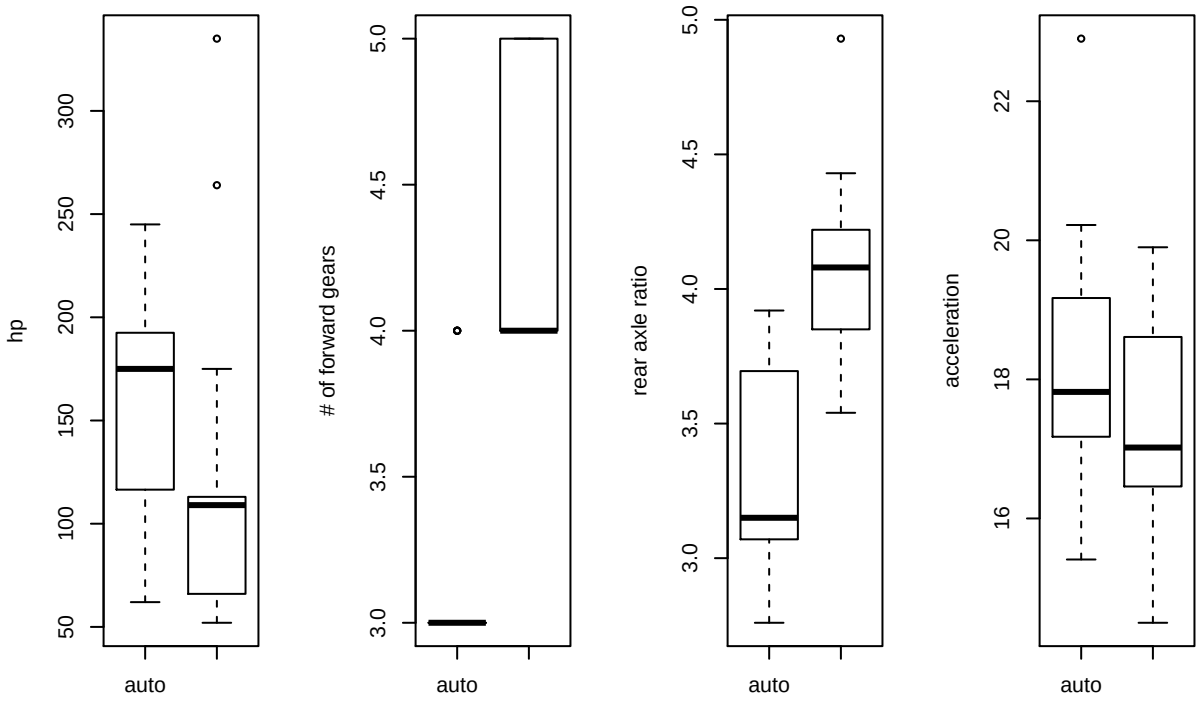
```r
t.test(wt ~ am, data = mtcars)
```

```
##
##  Welch Two Sample t-test
##
## data:  wt by am
## t = 5.4939, df = 29.234, p-value = 6.272e-06
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.8525632 1.8632262
## sample estimates:
##   mean in group auto mean in group manual
##             3.768895             2.411000
```

Given that automatic and manual cars had similar horsepower and performance/accelaration, it appears that they were equipped with bigger engines in order to produce more torque (of two engines with equal horsepower, the larger one will produce more torque). This makes sense because, due to the technological limitations of automatic transmission design in the seventies, these old automatic cars had fewer forward gears and smaller rear axle ratios than manual cars (or even modern automatic cars) and thus required more torque to achieve similar performance/acceleration.

```
par(mfrow=c(1,4))

boxplot(hp ~ am, data = mtcars, ylab="hp")
boxplot(gear ~ am, data = mtcars, ylab="# of forward gears")
boxplot(drat ~ am, data = mtcars, ylab="rear axle ratio")
boxplot(qsec ~ am, data = mtcars, ylab="acceleration")
```




```
t.test(hp ~ am, data = mtcars)
```

```
##
##  Welch Two Sample t-test
##
## data:  hp by am
## t = 1.2662, df = 18.715, p-value = 0.221
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -21.87858  88.71259
## sample estimates:
##   mean in group auto mean in group manual
##             160.2632             126.8462
```

```
t.test(gear ~ am, data = mtcars)
```

```
##
##  Welch Two Sample t-test
##
## data:  gear by am
## t = -6.8995, df = 22.568, p-value = 5.462e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -1.5264836 -0.8216945
## sample estimates:
##   mean in group auto mean in group manual
##             3.210526             4.384615
```

```
t.test(drat ~ am, data = mtcars)
```

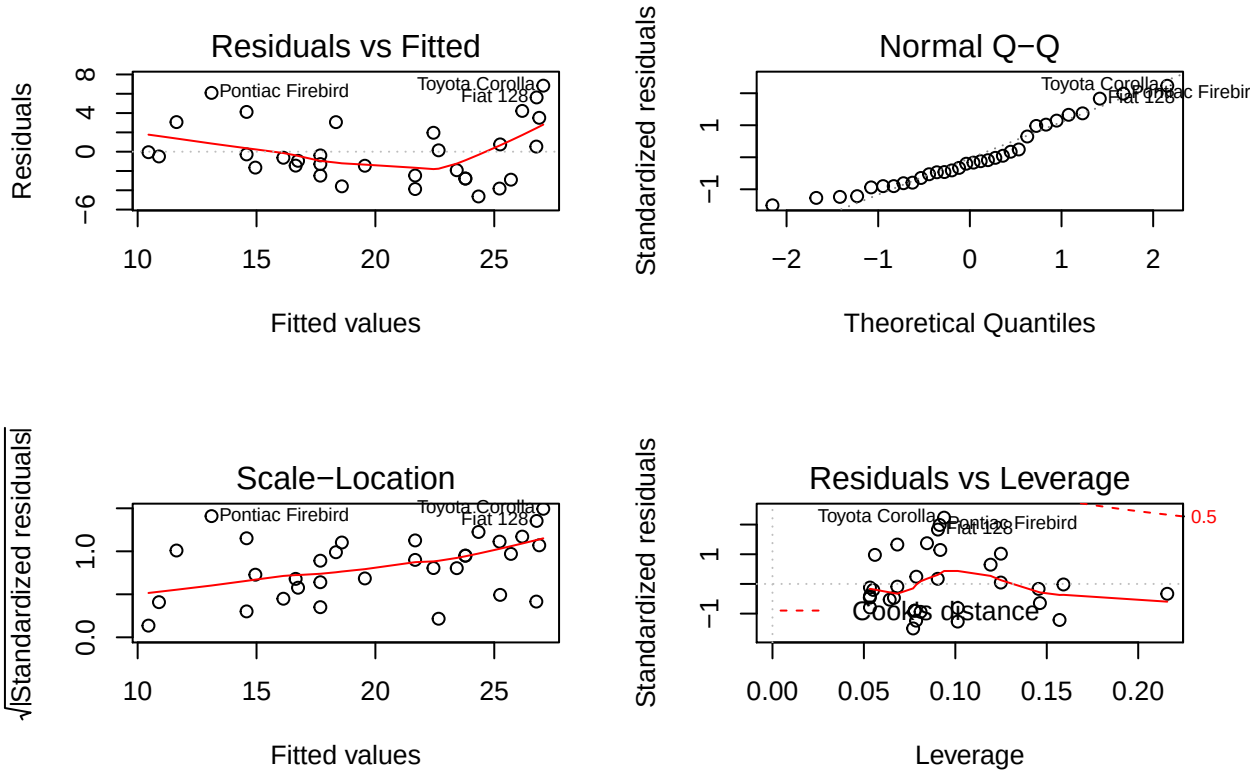
```
##
##  Welch Two Sample t-test
##
## data:  drat by am
## t = -5.6461, df = 27.198, p-value = 5.267e-06
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -1.0411183 -0.4862501
## sample estimates:
##   mean in group auto mean in group manual
##             3.286316             4.050000
```

```
t.test(qsec ~ am, data = mtcars)
```

```
##
##  Welch Two Sample t-test
##
## data:  qsec by am
## t = 1.2878, df = 25.534, p-value = 0.2093
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.4918522  2.1381679
## sample estimates:
##   mean in group auto mean in group manual
##             18.18316             17.36000
```

Model diagnostics

```
par(mfrow=c(2,2))
plot(fit2)
```


Residuals vs Fitted
Residuals
Fitted values
Pontiac Firebird
Toyota Corolla
Fiat 128
Normal Q−Q
Standardized residuals
Theoretical Quantiles
Scale−Location
√|Standardized residuals|
Residuals vs Leverage
Cook's distance
0.5
Leverage